Proceedings of 2005 IEEE Aerospace Conference, pp. 1-11, 2005.

Towards an Integrated Reasoner for Bearings Prognostics

Kai Goebel, Pierino Bonanni, Neil Eklund
GE Global Research
One Research Circle
Niskayuna, NY 12309
518-387-4194
{goebelk, bonanni, eklund}@crd.ge.com

Abstract—This paper describes an architecture and elements for an integrated prognostic bearings reasoner. The goal of the reasoner is to arrive at a reliable measure of damage accumulation, quantify its confidence, and aid in the assessment of remaining life. The reasoning integrates in-flight and post-flight functions. During flight, the tasks are primarily diagnostic and assess damage in real time using input from a plurality of sources, both sensor-based and model-based. The damage assessment is further refined after conclusion of the flight with full-order models and additional information from historical failure data, operational data, and inspection data. Load profiles from future missions are used to calculate the damage propagation which will allow the reasoner to assess remaining life. This paper will lay out the overall process and then focus in more detail on the in-flight reasoner. The operation of the architecture is demonstrated for bearing prognosis via an illustrative example.

TABLE OF CONTENTS

1. INTRODUCTION

Prognosis has been recognized as the "Achilles heel" [10] of condition-based maintenance (CBM). Here, the idea is to assess the condition of the system and perform maintenance when it is warranted. This is a departure from the traditional maintenance practice where components are managed to life limits based upon fleet wide statistics and average expected usage. The latter conservative approach requires the replacement of parts irrespective of how much of its useful life is actually expended. And indeed, most parts to be replaced under this philosophy still have a substantial amount of "life" left. If it were possible to account for the variability inherent to part manufacturing and operation, it should also be possible to change the life-limited replacement strategy to a condition-based parts replacement strategy [1], thus resulting in reduced cost of ownership with the same safety margin. If one could furthermore estimate the remaining life of a component, the whole paradigm of fleet management could be changed because it would be possible to not only perform maintenance at a convenient place and time taking into account variables such as part and staff availability, shop loading, etc. but also to plan more reliably future missions. To that end, a DARPA-sponsored program (of which the work reported herein is a part) addresses engine prognosis using advanced physics-based models, state-awareness sensors, and a prognostic reasoner to compute component capability and to quantify prediction-related variability and to provide system-wide capability assessment [2].

The physics-based models deal with mechanisms governing incipient damage at the material level, factoring in both full finite-element and reduced-order formats. The state awareness sensors measure material and system damage state, identify engine operation conditions, and update model predictions with advanced signal-acquisition and signal-conditioning methodologies. Finally, the prognostic reasoner fuses sensor and model-based information to assess residual component capability, calculate the uncertainty level for system predictions, and project a safe operational envelope for near-term engine usage [2].

This paper will focus on the prognostic reasoner demonstrated on the bearing systems. During bearings operation, initially localized spalls can initiate that may grow and ultimately result in loss of function. Important factors affecting damage initiation and damage propagation are changes in bearing loads and environment. Lubrication, presence of material defects, surface degradation, and external contamination all factor in to the bearing

[1]
[2] IEEEAC paper #1309, Version 3, Updated December 10, 2004

environment. Subsurface fatigue cracks are induced at locations of peak shear stress, become surface-connected, and lead to eventual liberation of material. It is important to assess the microstructural evolution, environmental embrittlement, cyclic hardening, and residual stress to calculate the propagation of bearing damage. The current state is determined by feeding direct sensor data and indirect parameters computed from sensor data into an ensemble of diagnostic algorithms as a basis for input to the fault-evolution and life models [2]. The information sources that the reasoner will rely on may potentially be obtained at different times during or in between flights and have different prediction horizons. They arrive at their conclusion either by direct measurement, models supported by measurements, or are simply triggered by measurements.

The reasoner described in this paper is represented as a multi-layered architecture comprising pre-processing, analysis, and post-processing steps. These steps are partitioned into modules where each module performs supporting tasks for both information processing and uncertainty management. In particular, the pre-processing step is partitioned into a temporal module that addresses issues surrounding the different time scale and sampling frequency of the information sources. In addition, the pre-processing step addresses how to deal with the different reliabilities of the information sources. The analysis step is partitioned into strengthening and weakening modules. These modules aggregate information sources of the models' output, the evidential sensor information, and the supporting extraneous information. They also strengthen or discount the found estimate using first principle and domain knowledge. The post-processing step is structured into exception handling modules that refine the output by dealing with special case situations.

2. BACKGROUND

Information Fusion

Finding synergy in using different information sources to assess system states has a long tradition within the fields of multivariate statistics and pattern recognition. Recently, the field of information fusion, and more specifically multi-classifier fusion has been recognized as a research area in its own right. Fusing information for prognostic purposes is a fairly new endeavor and will likely lead to the development of new techniques that are specialized to perform related tasks. Generally, techniques that lend themselves for prognostics are those that can provide continuous output. That means some classification techniques are not as readily applicable in the prognostic context because they provide only categorical output which would need to be artificially converted back into the continuous domain. However, if the answer sought is within the resolution of the answer bins provided, this may be an acceptable approach after all. Another big player in this context is the uncertainty associated with any estimate. If the uncertainty bounds are very wide, it does not make sense to demand a high resolution output. Rather, the output resolution should be commensurate with the residual uncertainty. In light of that, tools that lend themselves readily for a prognostic application would be fuzzy logic (in particular TSK-type models that provide functional output), neural nets, wavelet neural nets, or hybrid neuro-fuzzy techniques such as ANFIS [3]. Categorical classifiers that deal well with uncertainty integration include for example, Dempster-Shafer models, as well as others.

Sensors for Bearing Prognostics

Within the context of bearing prognostics, the particular requirements of bearing wear and bearing failure need to be taken into account. Here, wear particles build up over time even under normal operating conditions [4], leading to gradual damage accumulation that will introduce a bias of the starting point when faults occur. Particles generated by normal wear differ from particles seen during "abnormal" conditions such as spalling. Wear particles can be characterized with respect to their quantity, size, composition, and morphology [5] and therefore it would be desirable to capture and analyze those particles. While it is in practice easy to collect and measure the quantity of particles generated by debris collecting devices that are located in the oil scavenge line, it is somewhat more difficult to assess their morphology in an on-line application. However, the oil debris monitor counts the particles in bins of varying size ranges from which average particle size per bin can be computed. In addition, the cumulative mass can be calculated.

In addition to oil debris information, vibration information can be important to assess the onset of bearing failure. Vibration analysis has been proven to be very useful in machinery failure analysis [6]. Features from various transforms such as Fourier, Hilbert, and Wavelets can be useful in detecting and categorizing incipient faults. Indeed, vibration is a very sensitive measure that allows the early detection of bearing faults. This comes at the price of sensitivity to environmental effects [7] which are sometimes difficult to quantify or correct. In an aircraft engine, and in particular in a military engine, these changes can be significant. In contrast, oil debris feature provide relatively more reliable failure information [7].

It is appropriate to aggregate vibration and oil debris information to take advantage of the benefits of both. The fusion of information from oil debris and vibration information, along with knowledge about system and machinery history can result in interactions that may improve the confidence about system condition [4]. Howard and Reintjes [8] describe the benefits of using several information sources for fault detection, and discuss oil debris and vibration for helicopter gearboxes in particular.

Byington et al. [4] describe a fusion technique that correlates the failure mode phenomena with appropriate features. Dempsey et al. [7] report on the use of fuzzy logic to integrate oil debris and vibration information for gearbox faults where the output was quasi-action recommendations such as "OK, inspect, shutdown".

Prognostic Fusion

Integrating the capabilities and performance of the individual sensor systems and models through probabilistic weighting tools such as Bayesian and Dempster-Shafer allows the incorporation of a priori knowledge which has the promise to boost the performance of the overall system. In the example described by Orsagh et al. [1], model-based techniques were used when no diagnostic indicators were present and information derived from sensors such as oil debris sensors and vibration sensors were used when failure indicators are detectable. An important performance improvement is seen in allocating the weights to the information sources dynamically depending on whether the component is considered early or late in its remaining useful life cycle.

Other aggregation techniques include neural net based fault detection algorithms using Principal Component Analysis (PCA) output of a Fast Fourier Transform [9], and dynamic wavelet neural networks [10]. Garga et al. [11] describe a hybrid reasoning approach that is capable of integrating domain knowledge with test and operational data from an industrial gearbox. In this approach explicit domain knowledge is expressed as a rule-base and used to train a feedforward neural network.

The reasoner described in this paper will extend these findings by integrating a full spall model into the aggregation scheme and in addition by breaking the reasoning task into different modes. The modes are defined by either in-flight, post-flight, and future mission capability. This will allow to focus on more diagnostic tasks with short-term prognostic emphasis in the in-flight and post-flight mode. The future mission capability mode is entirely tasked with assessing remaining life given a set of future missions described by load and flight envelope conditions. Below we will examine the different modes in more detail. The remainder of the paper will then describe the work to date on the in-flight reasoner.

3. Integrated Reasoner Operating Modes

As mentioned above, the prognostic reasoner considered here is really a set of reasoners that will operate at various times during and after the flight. Depending on the time during or after a mission, its tasks will vary from aggregation of damage information to supporting the calculation of a remaining life estimate.

In-Flight Mode

During the flight, there are a number of information sources such as information derived from sensors that inform about the presence of bearing damage (Figure 1). Specifically, this information encompasses features derived from accelerometers that measure and assess vibration. Furthermore, information from debris monitoring devices is also used as a sensor-based input to the reasoner. In addition, a spall propagation model will provide information about the size and rate of increase of spalls. This model will use triggers from the reasoner to initiate its operation. That is, it will be dormant in the absence of evidence of bearing damage and fleet-wide statistics on bearing fatigue are used for low-level damage accumulation. Therefore, the reasoner will initially have to reliably provide diagnostic information about bearing damage to the spall propagation model. Once bearing damage has been established and the spall propagation model has been triggered, it will also need to integrate the information from the spall propagation model with the vibration and debris information.

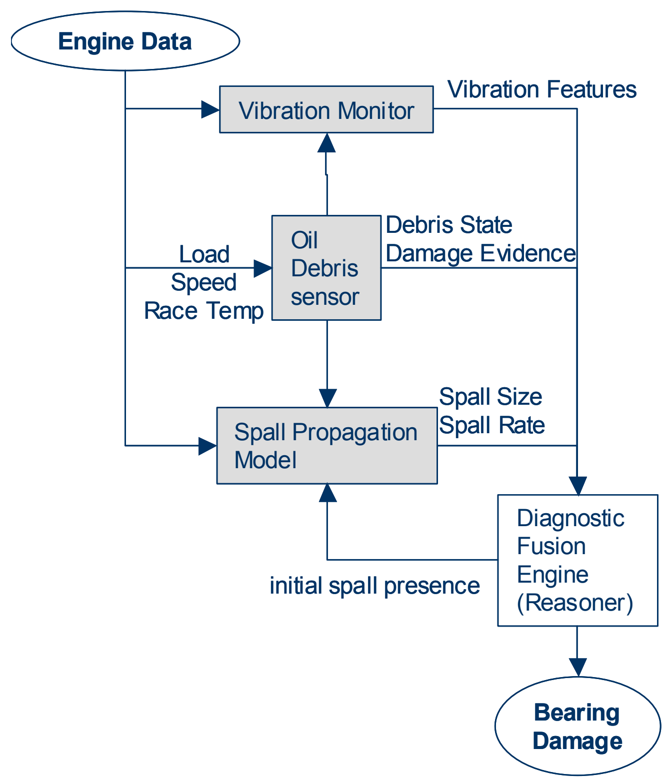


Figure 1 – In Flight Reasoner

Post-Flight Mode

During Post-Flight evaluation, the reasoner performs very similar tasks compared to the In-Flight tasks. However, while in-flight constraints may predicate reduced order models for the calculation of damage, the post-flight assessment allows for a full-order model to be run. It will rely on data that were collected during flight. In addition, there may be additional diagnostic information from maintenance, inspections, historical observations, etc. that can help to refine the diagnostic assessment about bearing damage. Figure 2 illustrates the post-flight scenario.

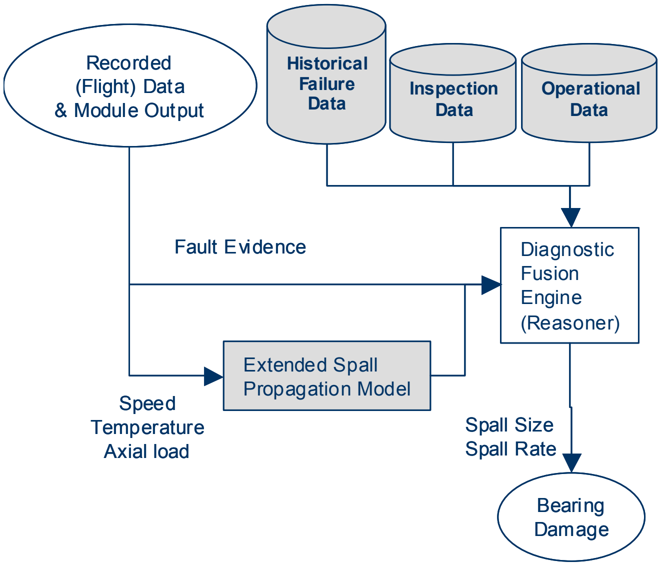


Figure 2 – Post-Flight Reasoner

Prognostic Mode

If a fault has been detected, the prognostic functions are executed on a set of future missions. Specifically, missions characterized in part by sequences of load, speed, and ambient conditions are used as input to the spall propagation model. In conjunction with the current damage state, the output of the spall propagation model will provide a damage profile into the future. This is illustrated in Figure 3.

The future mission prognostic reasoner has the form of a relatively simple decision logic that establishes whether an upper allowable damage limit has been surpassed for the given mission mix. There are different ways by which the reasoner can operate here based on user demand. In one instantiation, it will report both the profile of remaining life and information whether the envisioned missions can be completed without tripping the acceptable damage limit. In another instantiation, it will provide information back to the mission generation to prompt for additional mission runs when damage limits have not been reached. The goal of executing the damage propagation model with additional runs is to determine the damage propagation profile and to find the remaining life limit.

If no fault has been detected, the prognostic module gets de facto bypassed and is replaced by fleet statistics that are compiled on bearing fatigue data.

Figure 4 illustrates how the modules of the integrated bearing reasoner are connected. Note that during model development, the different modules are tuned off-line using raw data and ground truth data obtained during rig tests.

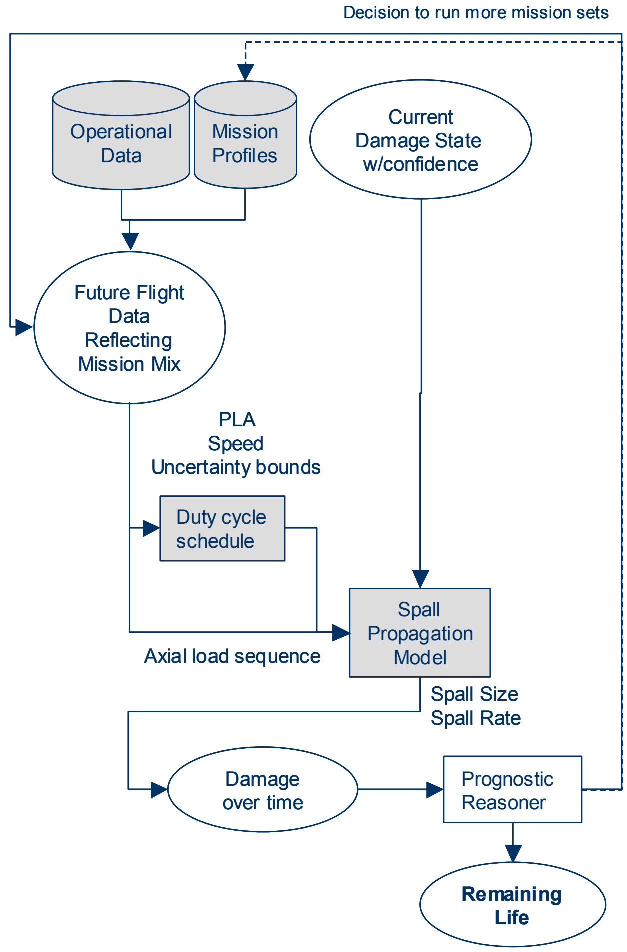


Figure 3 – Future Mission Prognostic Reasoner

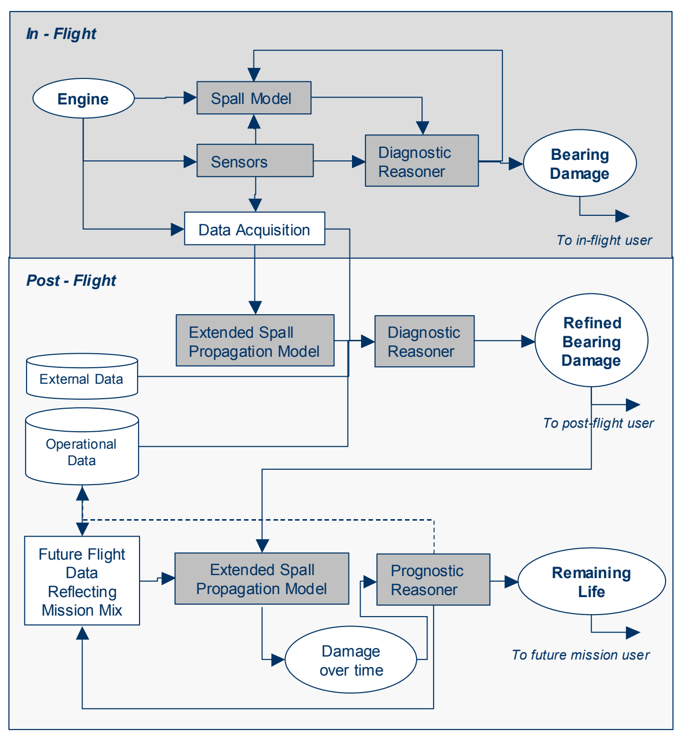


Figure 4 – Interactions of Integrated Bearing Reasoner Modules

4. IN-FLIGHT REASONER MODULES

The primary goal of the in-flight reasoner is to aggregate information from several sources to provide more reliable diagnostic and prognostic estimates than an estimate from a single source alone. Figure 5 shows the scheme of the proposed in-flight reasoner concept. It consists of two fundamental integration functions: integrating the outputs of individual damage models and evidential information and integrating domain knowledge. The two integration functions operate independently. While the damage model output and the evidential information are fed to the reasoner as inputs, domain knowledge is coded inside the fusion algorithms. Such independent integration design provides users the flexibility of adding more information sources and to integrate more domain knowledge or tune existing heuristics.

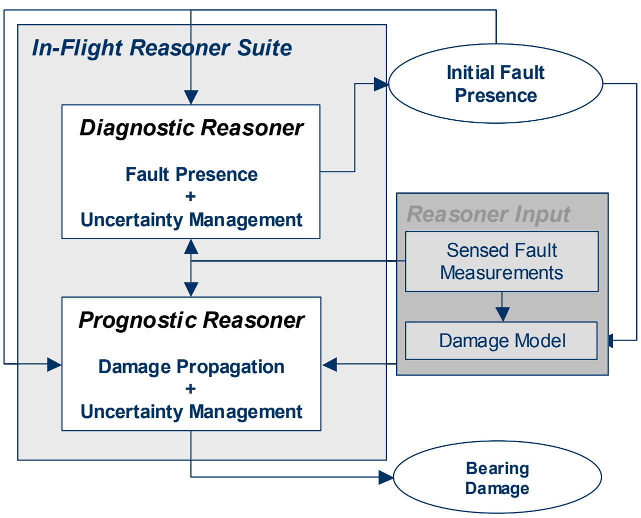


Figure 5 – Concept for In-Flight Reasoner

There are numerous approaches such as bagging and boosting [12], Dempster-Shafer [13], model-based approaches [14], fuzzy fusion [15] or statistics based approaches [16] that attempt to address the core aggregation functions. However, it has to be realized that the aggregation itself is only one function of the overall reasoner. In addition to combining information, it has to be ensured that the information that is being used provides the maximum information content. There are a number of issues that need to be dealt with prior to the actual aggregation. Specifically, the information needs to be checked for consistency, and it needs to be cleaned of outliers, noise, faulty or otherwise bad sensor information, it needs to be conditioned and formatted to allow a proper comparison. In addition to that, special cases need to be taken into account which, depending on the situation, should be done either before or after the actual aggregation step. To assist in these tasks, we suggest employing a sequential and parallel multi-layered configurations strategy. Elements from this configuration strategy have been proven successful in diagnostic fusion environments within project IMATE [17]. There, a hierarchical, multi-layer architecture [18] was demonstrated that implemented some of these concepts. Information from various diagnostic models and evidential information sources was combined and manipulated through a series of steps that increased and decreased the weight given to the information sources according to the strategies implemented in the respective layers of the fusion process.

In the following section we describe algorithmic concepts of the in-flight prognostic reasoner. In contrast to a diagnostic reasoner that has the task to determine the presence of a fault and therefore has as its output the fault category and perhaps an associated confidence, the in-flight prognostic reasoner needs to assess the presence of an initial fault condition and to report on the overall damage level plus an associated uncertainty. The most fundamental difference is in the second task, namely producing a damage assessment output that is in continuous format. This means that different aggregation techniques will need to be employed. Fundamentally, the reasoner functions are bundled in three major components (Figure 6): preprocessing, analysis, and post-processing. The preprocessing carries the majority of the burden for conditioning the information to be aggregated. The actual aggregation is performed in the analysis component. Finally, special cases are being addressed in the post-processing component.

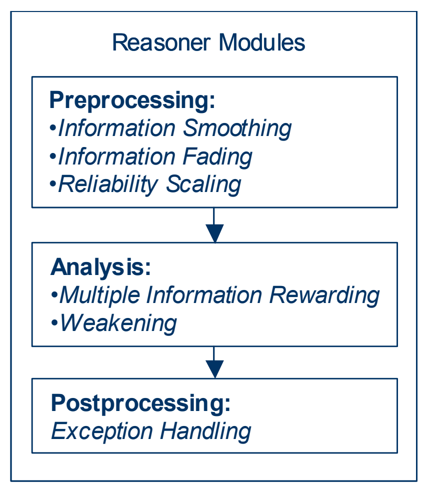


Figure 6 – Reasoner Modules

Major Components: Preprocessing

The Preprocessing component deals with manipulations of the damage model and the evidential information sources before the first fusion is performed. Here, temporal issues such as information source disagreement in time as well as integrating a priori reliability information are carried out. In addition, the preprocessing needs to concern itself with taking care of dubious sensor information (e.g., due to sensor failure, saturation), or process induced undesired sensor information (e.g., non-linear readings of cumulative debris sensors). We discuss these issues below before suggesting how to take care of them.

- *Nonlinearities* need to be taken into account to avoid illogical predictions. Nonlinearities can occur due to sensor noise but also, for example, when debris arrives at the debris monitor in irregular intervals. This nonlinear behavior should not necessarily translate into nonlinear damage accumulation.
- *Noise and outliers* can be observed in almost any sensor-based system. It is important to weed out outliers and to make the approach robust against noise while at the same time preserving the sensitivity to actual system changes.
- *Sensor failures* are a concern when the system depends on proper input. Recognizing sensor failure is typically done in a validation [19] step with a subsequent accommodation step
- *Saturation* can be considered a special case of sensor failure although the sensor is not technically failed. A sensors footprint may be limited to a certain range. While the sensor should be laid out to be able to operate properly in within the parameters of the system, this is a problem that may occur and against which a solution needs to be found. The accommodation of sensor saturation may be similar to the accommodation of sensor failure.
- *Disjoint response to event* The various sensors and sensor driven models may respond or indicate the event of interest at different times. The reasoner needs to be able to take these disjoint responses into account and translate them into a coherent output. A quick response needs to be balance with the need for confirmatory information.

The filter module sorts out outliers and noise and resolves temporary disagreement. Supporting functions employ Information smoothing and Information fading operators that remove outliers and trade off conflicting information.

Information smoothing—The challenge in dynamic systems is in providing a different reaction to situations where information sources agree and situations where information sources disagree. When information sources agree, and in the absence of evidence to the contrary, we postulate there is no reason for the fusion agent to change the collective opinion within that time step (there is still a minute chance of joint incorrect estimates). However, if information sources disagree, the fusion main module has to decide whether one source is correct and the other is not (and which) or whether an event has occurred that one source had no opportunity to see (for example because its update rate prevented a timely update). To help in this situation, we try to support the fusion main module by removing outliers and generally by smoothing information of individual sources in situations of agreement and by allowing quick updates when a changed event is indicated. This necessitates the need to provide a system status recognition tool that allows the two different strategies to work side by side.

The concept of information smoothing can be implemented via an exponential averaging time series filter with adaptive smoothing parameter [20].

Changes of the smoothing parameter will allow weeding out noise and outliers when no fault has occurred but reacting quickly to changes from one event to another [21].

Information Fading—Besides reducing variations in information during information agreement, we still need a mechanism to deal with the disagreement where one information source supports status s_1 at time t_1 and another information source comes to a different conclusion s_2 at a later time t_2. It is then necessary to account for the fact that s_2 may have occurred between t_1 and t_2. We postulate that the later status update needs to be given more weight in case of information disagreement to account for the possibility of occurrence of event s_2. The question is how much more weight s_2 (or how much less weight s_1) should have. We further propose that the discounting is a function of time passed between s_1 at time t_1 and s_2 at time t_2. The reason is that there is a higher possibility for event s_2 to occur when the period t_2-t_1 is larger and vice versa. In addition, the information sources must have information about their a priori performance. We propose to change the forgetting factor as the confidence value increases [22]. The idea of information fading is to discounting information as it ages when information sources disagree at different times (and no new update of both information sources can be obtained). We force the older information to "fade" as a function of time passed.

Reliability Scaling—The scaling module implements functionality allowing an information source that is a more reliable indicator for a particular fault to be weighted more heavily than an information source with a worse track record for that same fault [23]. Clearly, there should be higher trust in a more reliable information source, even when a less reliable information source indicates a strong confidence in its own opinion. However, even the more reliable information source is occasionally incorrect, and when the less reliable information source happens to be correct, it would be a mistake to always use the more reliable information source's output exclusively. Then again, it is not known when a particular information source is correct and when it is not. Therefore, we propose to use each information source's information scaled by the degree of its reliability. This reliability is encoded as a priori information in the confusion matrix. More specifically, it is the diagonal entries of the confusion matrix which we will use to weigh each classification output.

The concept of confusion matrices is firmly established in the realm of classification. Specifically, it allows the summary view of type I and type II errors. Table 1 shows an example confusion matrix.

Table 1 – Confusion Matrix

	Estimated no fault	Estimated fault
Actual no fault	TP	FP
Actual fault	FN	TN

Confusion matrix entries represent one operating point that embodies a particular tradeoff between false positives and false negatives. Whereas in diagnostics the aim is to classify a fault or precursor of a fault, a prognostics problem tries to make a judgment about the remaining life of a component. This has repercussions for the performance criteria used to measure the goodness of a tool and confusion matrices have not typically been used for prognostic evaluation. We will argue in the following that prognostic confusion matrices can be established. While it would seem a reasonable assumption to assess the performance by whether the estimate was on target or not, there will rarely be an estimate that is completely on the mark. However, this is in most cases not required anyhow. The question then is what is the acceptable tolerance for the problem at hand. We need to keep in mind that the utility of the error is oftentimes not symmetric with respect to zero (where the error is defined as the difference between actual remaining life and estimated remaining life). For instance, if the prediction is too early, the resulting early alarm forces more lead-time than needed to verify the potential for failure, monitor the various process variables, and perform a corrective action. On the other hand, if the failure is predicted too late, it means that this error reduces the time available to assess the situation and take a corrective action. The situation deteriorates completely when the failure occurs before a prediction is made that advises of critical system state. Therefore, given the same error size, it is in most situations preferable to have a positive bias (early prediction), rather than a negative one (late prediction). Of course, one needs to define a limit on how early a prediction can be and still be useful.

Therefore, two different boundaries for the maximum acceptable late prediction and the maximum acceptable early one can be established. Any prediction outside of the boundaries will be considered either a false positive or a false negative.

We define the prediction error [24] as

$$E(t) = [Actual\ \ time\ \ to\ \ failure\ \ (t) - Predicted\ \ time\ \ to\ \ failure\ \ (t)] \quad (6)$$

and we will report prediction results in terms of a histogram of the prediction error E(t). In particular, focus will be on two instances of $E(t)$:

- $E(t_r)$ - prediction error at the time when the critical zone (for example, within the next mission) is reached, and
- $E(t_0)$ - prediction error at the time when the failure occurs.

Incorrect classifications are typically classified as false negatives (FN) and false positive (FP). In the context of late or early predictions, these categorizations are based on the magnitude of deviation from true time of failure. Therefore, we will define the following limits as the maximum allowed deviations from the origin:

False Negatives A prediction is considered a false negative if one fails to correctly predict a failure more than t_{fn} time units later than the actual time to failure, i.e., $E(t_r) < -t_{fn}$ time units. Note that a prediction that is late more than t_r time units is equivalent to not making any prediction and having the failure occurring.

False Positives A prediction is considered a false positive if we fail to correctly predict a failure if the prediction is more than t_{fp} time units earlier than the actual time to failure, i.e., $E(t_r) > t_{fp}$ time units. We consider this to be excessive lead time, which may lead to unnecessary corrections.

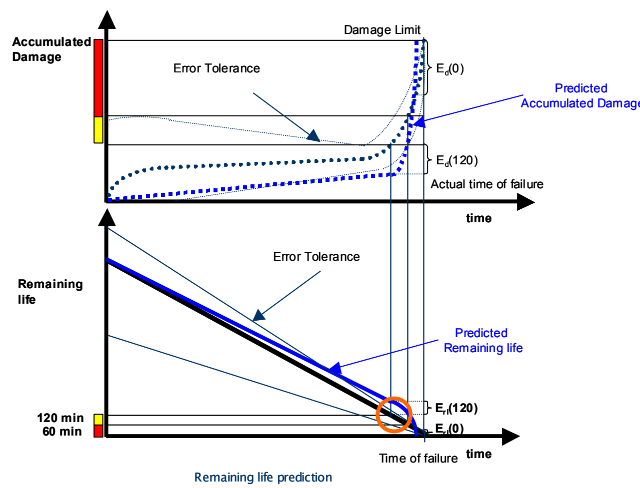


Figure 7 – Relationship between Accumulated Damage and Remaining Life

Figure 7 shows how the FP and FN are identified for a given prediction. It also shows the relationship between remaining life and accumulated damage. Figure 8 is a close-up of the relationship of time to failure and remaining life prediction. The graph also shows false positives and false negatives. These upper and lower limits are subjective and dependent on the problem domain

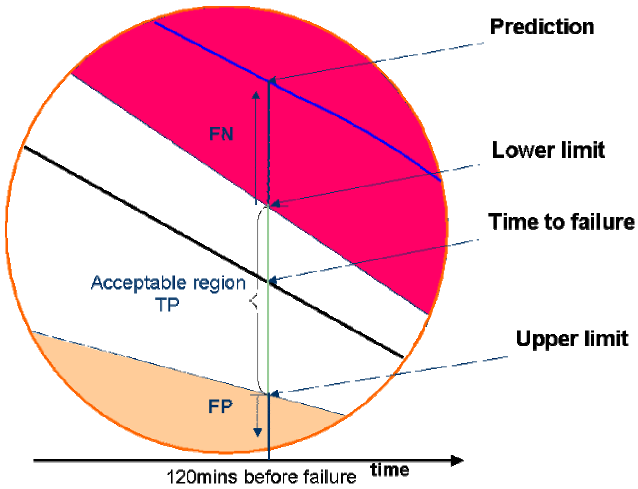


Figure 8 – False Positives and False Negatives in a Prognostic Context

Using the definitions of false positives and false negatives as outlined above, one can then choose a desirable operating point for prognostics. It also allows the design of the confusion matrix which is required for the reliability scaling.

Major Components: Analysis

Referring back to Figure 6, the analysis module performs the actual fusion. The functions contain both generic fusion strategies as well as domain knowledge heuristics. We have tested a number of different fusion techniques including weighted averaging and adaptive neuro-fuzzy inference systems. The latter has the advantage of automated learning capability while the former relies on the user to provide the appropriate weights. The two approaches mentioned both arrived at satisfactory results. Ultimately, the final fusion strategy needs to also include a provision to aggregate the uncertainty. For this purpose, confidence prediction neural networks [24] as well as typicality based approaches [25] and Dempster-Shafer-based regression methods [26] will be evaluated.

Multiple-Information Rewarding—State assessments expressed by different information sources which all agree should lead to a more confident assessment of the system state. This is the trivial case where coinciding opinions are rewarded. The rewarding scheme is accomplished by calculating a joint confidence value using for example a T-norm operator on the individual information source confidences that are in agreement.

Weakening—The weakening module integrates the fault information from similar equipment operating in tandem with the equipment under investigation (such as the bearing on the other end of the shaft). The hypothesis is that it is most unlikely that the same fault occurs on both systems at the same time. So, if the two systems issue the fault status at the same time, most likely, the condition is not fault related. Rather, it is due to a common environmental condition. Hence, the fault status should be discounted [27].

Major Components: Post Processing

Layers in this component deal primarily with exception handling.

5. Example Results

To demonstrate some of the concepts introduced here, we will present an illustrative example. Data for this example were simulated using anticipated sensor behavior as a guide. As the development of the different modules and the test rig experiments advance, the simulated data will be replaced (or augmented) by real data. Here, we show data of simplistic vibration sensors, debris collecting devices, and the spall propagation model. Random noise was added to all sensors where the vibration sensor had mostly amplitude varying noise with zero mean while the debris monitor also was exposed to random walk noise that produced local non-monotonic behavior. In addition, the debris sensor was exposed to a saturation type sensor fault which forced its upper value not to exceed a certain threshold. The fault event was triggered at different times for the different sensors but within a window of proximity. While the sensor behavior may not be truly reflective of the sensor behavior seen in the rig tests, this will allow us to test some of the concepts introduced earlier.

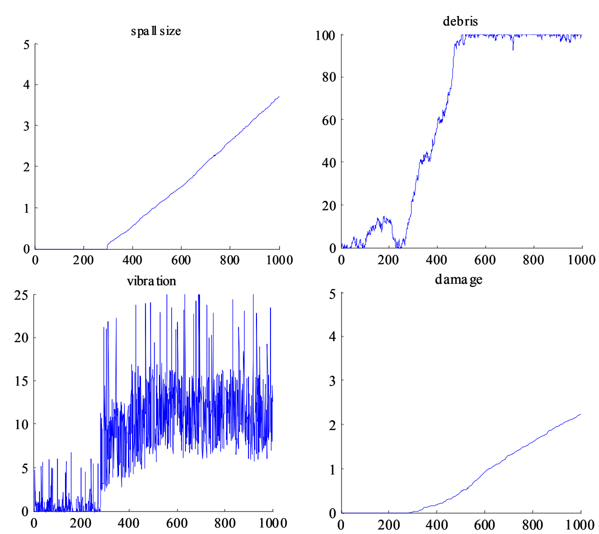


Figure 10 – Notional Sensor Output and Fused Output

Figure 10 shows output from the different sensor modules and the spall propagation model as well as the aggregated output information over time. Initially, no fault is present and the vibration sensor (Figure 11 (c)) is shown to be low on average with some amplitude spikes due to changes in operating conditions. However, at about time step 300, the vibration activity picks up considerably and increases in average amplitude before leveling out. During the no-fault condition, the primary task of the reasoner is to make sure that the spall size model is not triggered inadvertently due to any of the spikes. The debris monitoring sensor (Figure 11 (b)) shows some activity initially with low-order noise and a noticeable drop of its count before picking up again, Here, it is important for the reasoner to make sure that only monotonic information is being processed since the damage cannot have a negative slope. The debris sensor is also shown to saturate at a particular value. Again it is important for the reasoner to take this sensor behavior into account when aggregating information from the debris sensor. Figure 11 (a) shows the spall propagation model that has been correctly triggered at about time step 300 as a result of both the debris and the vibration information. The aggregated output is shown in Figure 11 (d) which shows the increasing damage of the bearing as a function of spall propagation model, vibration sensor, and debris sensor.

Data from rig tests will be included in the analysis throughout the rig tests. As of writing this report, they have not commenced yet. Results of an integrated engine level demonstration test will be reported in an upcoming paper.

6. Summary & Conclusions

We have presented initial concepts for an in-flight bearing prognostics reasoner. This reasoner is part of a suite of tools that comprise the integrated bearing prognostics system where the diagnostic information is aggregated during flight to provide short-term predictive capabilities. Longer term predictive capabilities are provided off-board and involve the calculation of component damage conditional on expected load conditions. The in-flight reasoner shown here allows the aggregation of both model-based damage information with sensor-based evidential information. Emphasis is placed on processing the information to ensure that the model – which relies on the reasoner to provide it with trigger information to get started – does not get kicked off either prematurely or too late. In addition, the sensor information and the model information are used within the reasoner jointly as mutual safeguards that provide important checks on their respective performance. Reasoner functions deal with conditioning the sensor signals as well as performing the aggregation. The former is embodied in several modules that modify the signal and manipulate an associated uncertainty. Different techniques (including T-norm based operators and ANFIS) are in place for the aggregation and have been tested on simulated data. The final choice will be driven by the best achieved performance. This decision will be made at the conclusion of a data collection period on rig tests which will also provide ground truth data.

Acknowledgments

This Material is based upon work supported by the Defense Advanced Research Projects Agency, Defense Sciences Office (DSO), Engine System Prognosis, ARPA Order No. Q517/00, Issued by DARPA/CMO under Contract Nos. HR0011-04-C-0001.

References

[1] R. Orsagh, J. Sheldon, and C. Klenke, "Prognostics/ Diagnostics for Gas Turbine Engine Bearings", Proceedings of ASME Turbo Expo 2003, Power for Land, Sea and Air, June 16 – 19, 2003, Atlanta, GA. GT2003-38075.

[2] J. Littles and M. Buczek, "Engine System Prognosis", Proceedings of Materials Science & Technology 2004, September 26-29, 2004, New Orleans, Louisiana, 2004.

[3] P. Bonissone and K. Goebel, "When will it break? A Hybrid Soft Computing Model to Predict Time-to-break Margins in Paper Machines", Proceedings of SPIE 47th Annual Meeting, International Symposium on Optical Science and Technology, Vol. #4787, pp. 53-64, 2002.

[4] C. Byington, T. Merdes, and J. Kozlowski, "Fusion Techniques for Vibration and Oil Debris/Quality in Gearbox Failure Testing", Proceedings of the International Conference on Condition Monitoring '99, pp. 113 – 118, 1999.

[5] B. Roylance and S. Raadnui, "The morphological Attributes of Wear Particles – Their Role in Identifying Wear Mechanisms", Wear, v 175, No 1-2, pp. 115-214, June 1994.

[6] M. Brennan, M. Chen, and A. Reynolds, A.G., "Use of Vibration Measurements to Detect local Tooth Defects in Gears", S V Sound and Vibration, v 31, no 11, pp. 12-17, Nov. 1997.

[7] P. Dempsey, R. Handschuh, and A. Afjeh, "Spiral Bevel Gear Damage Detection Using Decision Fusion Analysis", Proceedings of the Fifth International Conference on Information Fusion, Fusion 2002, vol. 1, pp. 94-100, 2002.

[8] P. Howard, and J. Reintjes, "A Straw Man for the Integration of Vibration and Oil Debris Technologies". Helicopter Health and Usage Monitoring Systems Workshop, G. Forsyth, ed., Defense Science and Technology Organisation Geeneral Document 197, no.1, pp. 131-136, 1999.

[9] R. Xu, M. Stevenson, C.-M. Kwan, and L. Haynes. "Bearing Monitoring", Proceedings of SPIE - Component and Systems Diagnostics, Prognosis, and Health Management, P. Willett, T. Kirubarajan, (Eds), vol. 4389, pp. 216-218, July 2001.

[10] G. Vachtsevanos, P. Wang, :Fault Prognosis Using Dynamic Wavelet Neural Networks", AUTOTESTCON Proceedings, IEEE Systems Readiness Technology Conference , 20-23 Aug. 2001, pp. 857 – 870, 2001.

[11] A. Garga, K. McClintic, R. Campbell, C.-C. Yang; M. Lebold, T. Hay, C. Byington, "Hybrid Reasoning for Prognostic Learning in CBM Systems", IEEE Proceedings Aerospace Conference, vol. 6, 10-17 March 2001, pp. 2957 – 2969, 2001.

[12] Y. Freund and R. Schapire, "A Short Introduction to Boosting", J. Japanese Society Artificial Intelligence, Vol. 14, No.5, pp. 771-780, 1999.

[13] P. Smets, "What is Dempster-Shafer's model?" Advances in the Dempster-Shafer Theory of Evidence, Yager, R., Fedrizzi, M., and Kacprzyk, J., (Eds.), John Wiley & Sons, New York, pp. 5-34, 1994.

[14] M. Nelson and K. Mason, "A Model-Based Approach to Information Fusion". Proc. Information, Decision, and Control, pp. 395-400, 1999.

[15] A. Loskiewicz-Buczak, R. Uhrig, "Decision Fusion by Fuzzy Set Operations", Proc. third IEEE Conf. Fuzzy Systems, Vol. 2, pp.1412-1417, 1994.

[16] N. S. V. Rao, "Finite Sample Performance Guarantees of Fusers for Function Estimators, Information Fusion, Vol. 1, no. 1, pp. 35-44, 2000.

[17] M. Ashby and W. Scheuren, "Intelligent Maintenance Advisor for Turbine Engines (IMATE)", Proceedings of the IEEE Aerospace Conference, 11.0309, 2000.

[18] K. Goebel, "Architecture and Design of a Diagnostic Information Fusion Tool", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 15 (4), Sept. 2001, pp. 335-348, 2001.

[19] K. Goebel, A. Agogino and S. Alag, "Probabilistic and Fuzzy Methods for Sensor Validation and Fusion in Vehicle Guidance: A Comparison", Proceedings of ISATA '97, 30th International Symposium on Automotive Technology & Automation; Florence, Italy, Vol. 1, pp. 711-719, 1997.

[20] K. Goebel, "Decision Forgetting and Decision Smoothing for Diagnostic Information Fusion in Systems with Redundant Information", Proc. SPIE: Sensor Fusion: Architectures, Algorithms, and Applications IV, Vol. 4051, pp. 438-445, 2000.

[21] P. Khedkar, and S. Keshav, "Fuzzy Prediction of Time Series", Proceedings of the IEEE International Conference on Fuzzy Systems, San Diego, 1992.

[22] K. Goebel, M. Krok, and H. Sutherland, "Diagnostic Information Fusion: Requirements Flowdown and Interface Issues", Proceedings of the IEEE 2000 Aerospace Conference – Advanced Reasoner and Information Fusion Techniques, vol. 6, pp. 155-162, 2000.

[23] K. Goebel and S. Mysore, "Taking Advantage of Misclassifications to Boost Classification Rate in Decision Fusion", Proceedings of SPIE, Sensor Fusion: Architectures, Algorithms, and Applications V, 2001, pp. 11-20, 2001.

[24] P. Bonissone and K. Goebel, "When will it break? A Hybrid Soft Computing Model to Predict Time-to-break Margins in Paper Machines", Proceedings of SPIE 47th Annual Meeting, International Symposium on Optical Science and Technology, Vol. #4787, pp. 53-64, 2002.

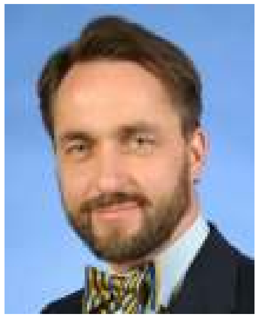
[24] G. Vachtsevanos, "Performance Metrics for Fault Prognosis of Complex Systems", Proceedings of AUTOTESTCON 2003, pp. 341-345, 2003.

[25] D. Dubois, H. Prade, and J. Rossazza, "Vagueness, Typicality, and Uncertainty in Class Hierarchies", International Journal of Intelligent Systems, vol. 6, no. 2, Fuzzy Methods for Knowledge-Based Systems, p 167-183, 1991.

[26] S. Petit-Renaud and T. Denoeux, "Nonparametric Regression Analysis of Uncertain and Imprecise Data Using Belief Functions", International Journal of Approximate Reasoning, Vol. 35, No. 1, pp. 1-28, 2004.

[27] K. Goebel and W. Yan, "Fusing Binary and Continuous Output of Multiple Classifiers", Proceedings of the Fifth International Conference on Information Fusion, Fusion 2002, vol. 1, pp. 380-387, 2002.

BIOGRAPHY



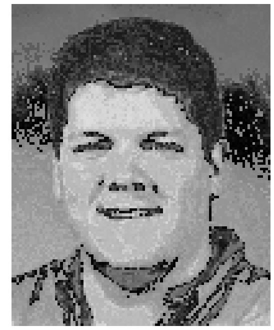
Kai Goebel *received the degree of Diplom-Ingenieur from the Technische Universität München, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively.*

Dr. Goebel joined General Electric's Corporate Research and Development facility in Schenectady, NY in 1997 as a computer scientist after working as a visiting postdoctoral fellow at UC Berkeley from 1996 to 1997. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He has fielded numerous applications for aircraft engines, transportation systems, medical systems, and manufacturing systems.

Dr. Goebel is an adjunct professor of the CS Department at Rensselaer Polytechnic Institute (RPI), Troy, NY, since 1998 where he teaches classes in Soft Computing.

Pierino Bonanni *received his S.B., S.M., E.E., and Ph.D. degrees in Electrical Engineering from the Massachusetts Institute of Technology in 1983, 1983, 1984, and 1991, respectively.*

Dr. Bonanni joined GE Global Research Center in Niskayuna, New York, in 1991. His specialization is in signal and image processing, with emphasis on detection, estimation, and inverse problems arising in controls and diagnostics applications. He has performed and led a variety of applied research efforts in fault detection. machine vision, robotics, automated manufacturing, satellite navigation and control, non-destructive inspection, and remote sensing. He is the recipient of several GE team and project management awards, and holds 7 U.S. patents.



Neil Eklund *received B.S. in 1991, two M.S. degrees in 1998, and a Ph. D. in 2002, all at the Rensselaer Polytechnic Institute.*

Dr. Eklund was a research scientist at the Lighting Research Center from 1993 to 1999. He was in the network planning department of PSINet from 1999 to 2002, before joining General Electric Global Research in Niskayuna, NY in 2002. He has worked on a wide variety of research projects, including early detection of cataract using intraocular photoluminescence, multiobjective bond portfolio optimization, paint inspection lighting optimization, multiobjective optimization of permanent virtual circuits for a frame relay service, and the influence of lighting on office worker performance. His current research interests include color measurement, color and spatial vision, genetic algorithms for multiobjective optimization, and applied soft computing.